

Amplitude Modulation and Loudness in Cochlear Implantees

COLETTE M. MCKAY^{1,2} AND KATHERINE R. HENSHALL¹

¹*Department of Otolaryngology, The University of Melbourne, Melbourne, 3054, Australia*

²*School of Psychological Sciences, The University of Manchester, Manchester, M13 9PL, UK*

Received: 10 May 2009; Accepted: 9 September 2009; Online publication: 2 October 2009

ABSTRACT

The effect of amplitude modulation of pulse trains on the loudness perceived by cochlear implantees was investigated for different overall levels of the signal, modulation depth and the carrier rate of the pulse train. Equally loud and threshold levels were determined for a variety of signal levels, modulation depths and carrier rates in six cochlear implantees. The pattern of results was consistent with the predictions of a previously published loudness model of McKay et al. (*J Acoust Soc Am* 113:2054–2063, 2003). The degree to which the loudness of modulated stimuli differed from the loudness elicited by an unmodulated pulse train with equivalent average current depended on the modulation depth and the absolute current level of the unmodulated stimulus. The effect of carrier rate on this measure was predictable solely from the effect of rate on absolute current level for equal loudness. The results have important implications for the interpretation of experiments measuring modulation detection that do not control loudness cues. We show that several previously published results regarding the effect of carrier rate and added noise on modulation detection could be reinterpreted in the light of these findings.

Keywords: loudness model, electrical stimulation, temporal modulation

INTRODUCTION

Although the loudness of constant-current pulse trains has been studied by many researchers, the effect of amplitude modulation on loudness has not

received much attention. The aims of this study were to quantify the effect of amplitude modulation of pulse trains on loudness and to test whether the loudness model of McKay et al. (2003) could successfully predict the results. This study is important for two reasons: First, it tested the relevance of the McKay model to the control of loudness in speech processors; second, it provided important information for the interpretation and design of psychophysical studies using amplitude modulated stimuli.

Two studies (Zeng and Shannon 1995; Zhang and Zeng 1997) examined the effect on loudness of amplitude modulation of analog electrical stimuli. Both studies showed that the effect of modulation on loudness was level-dependent, with the root mean square current level determining loudness at levels near threshold and the peak current level determining loudness for loud sounds. These studies have not been replicated using pulsatile stimuli, which is surprising considering that pulsatile stimuli are almost universally used in commercial implant devices. Indeed, as discussed later, all studies that have examined modulation detection have implicitly assumed that modulating a pulsatile current around a mean produces no loudness change (i.e. the mean current determines the loudness).

Underlying the results of Zeng and Shannon and of Zhang and Zeng is the shape of the loudness versus current function. Studies using pulsatile stimulation and loudness estimation or matching techniques (Shannon 1985; Zeng and Shannon 1994; Chatterjee 1999) have variously fitted the loudness versus current data with either a power function or exponential function. The loudness model of McKay et al. (2003) was developed using data ranging across different rates of pulsatile stimulation (500–4,000 Hz) encompassing a wider current range. The data in that study were best fit by a log(loudness) versus log(current) function whose slope was linear for low absolute current levels (less than about 200 μA in MP mode) but greater than exponential for high absolute current levels.

Correspondence to: Colette M. McKay · School of Psychological Sciences · The University of Manchester · Manchester, M13 9PL, UK. Telephone: +44-161-3061671; fax: +44-161-2753373; email: Colette.mckay@manchester.ac.uk

There are two unique features of the McKay model. First, the loudness predictions are based on absolute current levels (relative to a single subject-dependent 'kneepoint' value as described below), rather than the level relative to the dynamic range of the particular stimulus. This leads to a prediction that the shape of the current versus loudness function will differ for stimuli of different rates. Second, the practical prediction method based on the loudness model allows the relative loudness of any complex temporal-spatial pattern of electrical pulses to be predicted, as it incorporates the effects of temporal and spatial loudness summation. The general applicability of the prediction method to signal processing has been verified by its use in a signal processing strategy called SPeL (McDermott et al. 2003; McDermott et al. 2004).

The experiments reported here investigated the effects of amplitude modulation on loudness in pulsatile stimuli of differing rates (500, 1,000 and 8,000 Hz) and using various modulation depths and modulation rates of 250 and 500 Hz. The McKay model was used to predict the pattern of results for the particular stimuli used. The range of modulation rates and carrier rates were chosen to capture the interesting effects predicted by the model. The study had ethical approval from the Human Research Ethics Committee of the Royal Victorian Eye and Ear Hospital.

METHODS

Subjects and stimuli

Six users of the CI24M cochlear implant manufactured by Cochlear Ltd participated in this study. All were adults who were post-lingually deafened. Their details are shown in Table 1. All were experienced in psychophysical procedures.

The stimuli used in the experiments were either constant-current pulse trains or amplitude-modulated pulse trains on single electrodes using monopolar (MP1+2) mode. The modulation was sinusoidal current modulation in cases where the carrier rate was more than double the modulation rate and alternated peak to

peak in cases where the carrier rate was twice the modulation rate. Phase durations were 26 μ s, with an interphase gap 8.4 μ s, and the stimuli duration was 500 ms. Carrier rates used were 8 and 1 kHz (with 500 Hz modulation) and 500 Hz with 250 Hz modulation. Note that these modulation rates are higher than those associated with temporal envelope cues in speech but were chosen because they lead to a prediction of a steady, not fluctuating, loudness and thus were the most appropriate to test the loudness model. A discussion of how the model would apply to lower frequency modulations is contained in the discussion. Modulation depths were designated in terms of clinical current steps between peak and valley and varied between 0 and 20 in steps of 5. In this device, each current step represents a 0.176-dB change in current. For each modulated stimulus, the peak level was fixed at either 90% or 60% of the dynamic range (expressed in current steps) of the analogous constant-current stimulus. For that reason, the average currents in the modulated stimuli differed for different modulation depths. Additionally, absolute threshold measures were taken with all stimuli. When adapting the level of a modulated stimulus in the procedures, all pulses in the stimulus were incremented or decremented by equal numbers of current steps.

Experimental procedures

The threshold and comfortably loud levels for each (unmodulated) carrier rate were first determined so that the peak levels of the modulated stimuli could be fixed at 60% and 90% of this dynamic range. The comfortably loud level was determined by a method of adjustment, whereby the current was slowly incremented until the subject indicated a 'too loud' percept and then decremented to a loud but comfortable point. Thresholds were determined using an adaptive four-interval four-alternative forced choice procedure in which the stimulus could occur in any of the four intervals. A two-down one-up rule asymptoting to 71% correct was used. Adaptive steps were four current steps for the first two turns and two current steps for the remaining eight turns, the last six of

TABLE 1

Details of subjects who participated in the experiments

<i>Subject</i>	<i>Aetiology</i>	<i>Duration of profound deafness (years)</i>	<i>Duration of implant use (years)</i>
S1	Unknown	12	2
S2	Hereditary progressive	2	2
S3	Hereditary progressive	2	5
S4	Chronic infection	2	5
S5	unknown	11	7
S6	Familial and chronic infection	35	5

All were adults who used the CI24M implant manufactured by Cochlear Ltd

which were averaged to determine the threshold. Two measures of threshold were obtained for each stimulus and the results averaged.

For each of the 24 modulated stimuli (four non-zero modulation depths, three carrier rates and two overall levels), the loudness was matched to a constant-current pulse train of equal rate. In the loudness balance procedure, the reference stimulus (which in alternate trials was the modulated and the unmodulated stimulus) was fixed and the stimulus to be balanced was adjusted in an adaptive two-interval, two-alternative forced choice task in which the subject had to respond with which interval contained the louder sound. The stimuli were presented in random order with an inter-stimulus interval of 500 ms. The test stimulus (starting at a different level in each run) was incremented or decremented according to each response. Adaptive steps were four current steps for the first two turns and two current steps for the remaining eight turns, the last six of which were averaged to determine the current for balanced loudness. A total of four balances (with unmodulated stimulus and modulated stimulus both being the reference twice) were performed. When the modulated stimulus was the reference (runs 1 and 3), its level was fixed using the peak current levels chosen above (60% and 90% of the dynamic range of the unmodulated stimulus). When the unmodulated stimulus was the reference (trials 2 and 4), its level was set to the balanced current level found in trial 1 or 3. The four measures of current difference (in current steps) between the peak current in the modulated pulse train and the current in the equally loud constant-current pulse train were averaged and noted for analysis.

In addition to the loudness-matching procedure above, absolute detection thresholds of all the modulated stimuli were determined using the same procedure as for the unmodulated stimuli. Analogous data as in the loudness balance procedure were determined (the difference between the peak current in the modulated stimulus and the fixed current in the unmodulated stimulus when both stimuli were at threshold).

The model predictions

To assist the reader, a brief description of the McKay loudness model is given here and a description of the method used to model the predictions in this experiment. The loudness model was developed to account for the temporal characteristics (interpulse intervals) within single-electrode stimuli (McKay and McDermott 1998), as well as the effects of spatial and temporal separation in multiple-electrode stimuli (McKay et al. 2001). Essentially, the *temporal effect* of interpulse interval can

be modelled by the effects of central temporal integration (a time window of several milliseconds) and peripheral neural refractoriness. The investigation of *spatial separation effects* in dual-electrode stimuli showed very little (none on average) effect of electrode separation, implying that the expected impact of refractoriness in the neural population activated by both electrodes (predicting reduced overall activity and hence loudness) was mitigated by some other factors (postulated to be a non-linear relationship between neural activity and loudness). As a consequence of the absence of effect of electrode separation, we made the simplifying assumption that each pulse in a pulse train that falls within the integration window contributes independently to the overall loudness, regardless of where the pulse is spatially located. This assumption led to a simple practical method of predicting the relative loudness of any complex pulsatile stimuli (McKay et al. 2003). This is the method that was used in the study described here to predict the loudness of the modulated relative to the unmodulated stimuli.

Essentially the loudness contribution of *each pulse* of a stimulus within the temporal integration window is given by a relationship of the form:

$$\log(L) = 0.019 \times x + 0.21 \times e^{(x-x_0)/7} + k \quad (1)$$

where L is loudness, x is the current of the pulse (in clinical current steps), k is an arbitrary constant and x_0 is a kneepoint current above which the exponential term becomes important. The kneepoint current is generally at a current level that is below but close to the maximum loudness for a low-rate stimulus (500 Hz or less) for that subject/electrode. The numerical values in Eq. 1 can vary between patients. The multiplier in the linear term is relatively constant across implantees (range of 0.015–0.021 approximately). The values in Eq. 1 are typical for the subjects in the 2003 derivation paper and a larger corpus of unpublished data. To determine the relative loudness of two electrical stimuli, a temporal window is first selected. To avoid the necessity of averaging the calculations across many such windows along the stimulus, the repetition period (if any) of the stimuli can be used as the summation window, so long as it is less than or equal to the temporal integration window of the model (7 ms equivalent rectangular duration ERD). Within the window, the loudness contributions of the n pulses are determined by Eq. 1, and the total loudness (L) is merely the sum of the individual loudness (L_n):

$$L(\text{total}) = \sum_n L_n \quad (2)$$

In modelling the expected results of the current experiment for a ‘virtual’ subject, we used the numerical values as in Eq. 1. The kneepoint current

step was arbitrarily chosen to be 200 for our virtual subject. We then determined the current levels to use for the different rates of stimulation by first selecting a maximum comfortable level of 209 for the 500 Hz stimulus (to be consistent with the choice of knee-point) and a perceptual dynamic range of loudness from $L=10$ (threshold) to $L=100$ (comfortable). The currents for different perceptual levels (loudness values) at different rates of stimulation can then be obtained using Eqs. 1 and 2.

Then, for each modulated stimulus, the overall loudness within the temporal integration window (2 or 4 ms) was computed, and the level of the corresponding non-modulated stimulus that produced the same loudness was determined. Note that the parameters in the loudness model are for a ‘virtual subject’ and were not determined specifically for each actual subject in this experiment, so the modelled results represent the general pattern of expected results rather than the exact data for each experimental subject.

RESULTS

The model predictions

The model predictions are shown in Figure 1. For reference, the figure also shows data predictions that would occur if the peak current in the modulated stimulus determined the loudness or the average current determined the loudness. It can be seen that the model predictions generally fall between the ‘peak’ and ‘average’ predictions and that the model predictions differ for different carrier rates. For the highest rate (8 kHz), the average current gives a fair approximation of the loudness prediction for all levels between threshold and 90% of the dynamic range. For the 500 Hz and 1 kHz carrier rates, threshold loudness was predicted to be close to the average current, but at higher current levels, the predictions move closer to the ‘peak’ current, and this effect is greater for the lower carrier rate.

The experimental data

The threshold and comfortably loud levels of the unmodulated carrier rates for the six subjects are shown in Table 2. These data show the well-known effect of rate on current levels for equal loudness.

The experimental data comparing loudness of modulated and unmodulated stimuli were averaged across subjects and plotted in the same way as the model data. The results are shown in Figure 2. It can be seen that the general pattern of model predictions (Fig. 1) matches the data pattern quite well, with data for all levels at the highest carrier rate and the

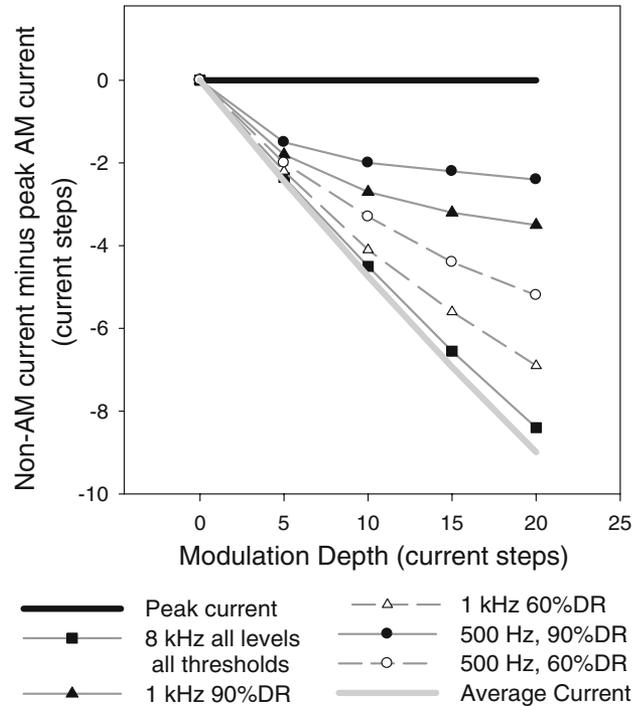


FIG. 1. Model predictions for the current difference between the fixed current of the non-modulated stimulus and the peak current in the modulated stimulus (in current steps). Also shown on the graph are the predictions if the peak current determined the loudness (*thick black horizontal line* at 0 difference) and if the average current determined the loudness (*thick diagonal grey line*). Note that the lines for five of the predictions (all threshold levels and all 8 kHz carrier rates) all fall very close to the ‘average current’ prediction line.

threshold data at the lower carrier rates, falling close to the average-current prediction line. In contrast, for lower carrier rates, the data moved towards the ‘peak’ prediction line for progressively higher stimulus levels but by differing amounts for different rates, with the greatest effect at the lowest rate. We tested the hypothesis that the average adjustment for all 8 kHz stimuli and the threshold stimuli was close to but less than that predicted by the average current for the modulation depth of 20 current steps. The average adjustment for these stimuli was -7.96 steps (with 95% confidence interval of -6.78 to -9.14). The prediction for ‘average current’ (-8.99 steps) lies just within the confidence interval. In contrast, the ‘average current’ prediction lies well outside the confidence intervals (min, max) for adjustment for the other stimuli: 1 kHz 60%DR (-1.7 , -6.2), 1 kHz 90%DR (-0.6 , -5.5), 500 60%DR (-0.6 , -6.7) and 500 Hz 90%DR (-0.6 , -1.5).

Equation 1 predicts that the effect of level on loudness is linear (on log–log scale) for low absolute current levels and becomes progressively non-linear at higher absolute current levels. That is, the non-linear effect is present not at higher *perceptual* levels but at high *absolute* current levels. At higher absolute levels, the higher level pulses within a modulated stimulus

TABLE 2

Thresholds, comfortably loud levels and dynamic range (all in current steps) for the three carrier rates (unmodulated) used in the experiment

Subject and rate	Carrier rate	Threshold	Comfortably loud	Dynamic range
S1	500	121.3	170	48.7
	1,000	105.8	164	58.2
	8,000	52.3	134	81.7
S2	500	116.8	189	72.2
	1,000	100.5	184	83.5
	8,000	49.7	160	110.3
S3	500	149.5	186	36.5
	1,000	132	180	48
	8,000	81.5	162	80.5
S4	500	132	176	44
	1,000	112.8	191	78.2
	8,000	64.4	163	98.6
S5	500	161.3	193	31.7
	1,000	152.7	192	39.3
	8,000	101.5	174	72.5
S6	500	169.3	206	36.7
	1,000	150.3	201	50.7
	8,000	105	163	58

contribute a greater proportion of the overall loudness than in the same modulation pattern at a lower overall level.

To test the effect of absolute level more directly in the data, both for the virtual subject and for the experimental subjects, the data or model predictions were first normalised to a metric that could be averaged across modulation depths: The current differences between the ‘peak’ and equally loud non-modulated stimuli were divided by the modulation depth. That is, the new data were the difference between the equally loud constant current and the peak level in the AM stimulus, expressed as a proportion of the modulation depth. These data, plotted against absolute peak current level for the virtual subject, are shown in Figure 3, and experimental data averaged across the experimental subjects are shown in Figure 4. It can be seen in Figure 3 that the absolute current determines the adjustment for equal loudness and that rate or level in the dynamic range only affects the prediction via their effect on the absolute current used. The data in Figure 3 are reproduced as the dashed line in Figure 4 for comparison with the experimental data. The dashed line is horizontally offset between the figures to account for difference between virtual and experimental subjects in the patient-specific kneepoints in Eq. 1. The kneepoint was arbitrarily set to 200 current steps for the virtual subject. However, the assumed kneepoints for the experimental subjects would be lower by a difference consistent with the differences in maximum current level for the 500 Hz stimulus

(209 for the virtual subject and an average of 182 for the experimental subjects). It can be seen from Figure 4 that the experimental data reproduce the same effect of absolute level as does the model. In fact, the fit of the model to the data is surprisingly good when one considers that the specific parameters used in Eq. 1 for the virtual subject are likely to be somewhat different from those that accurately describe each experimental subject’s loudness growth function.

DISCUSSION

General discussion

The results of this experiment have shown that the practical method proposed by McKay et al. (2003) to predict the relative loudness of different complex electrical stimuli is able to correctly predict, for modulation rates of 250 Hz or greater, the average effect of amplitude modulation of a pulse train on the

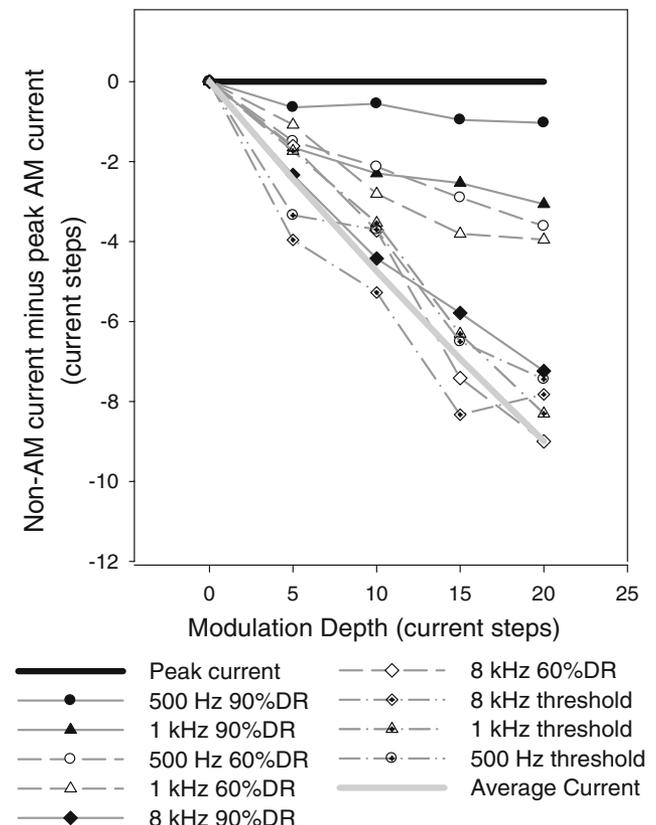


FIG. 2. Experimental data for the current difference between the fixed current of the non-modulated stimulus and the peak current in the modulated stimulus (in current steps). The same reference lines as in Figure 1 are plotted showing the predictions if the peak current determined the loudness (black horizontal line at 0 difference) or the average current determined the loudness (diagonal grey line). Error bars are omitted from the figure to avoid clutter, but the average standard deviation was 2.2 current steps (or average standard error of 0.9 current steps).

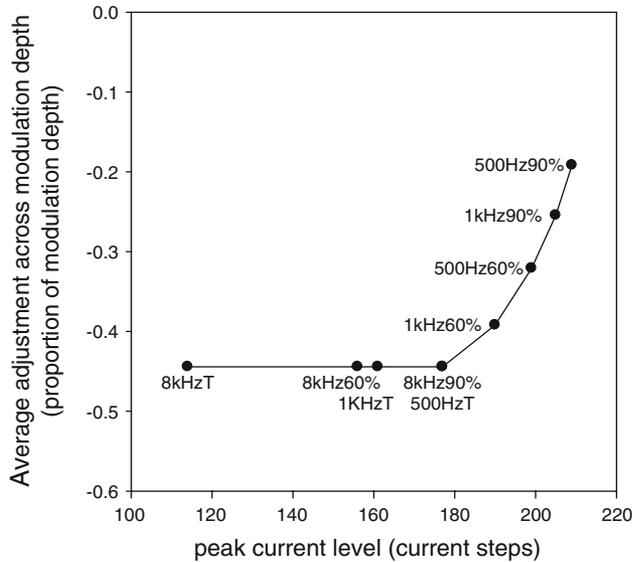


FIG. 3. Model predictions for the difference (averaged across modulation depth) between peak current in the modulated stimulus and the equally loud fixed current level, expressed as a proportion of the modulation depth. Labels are to the left or underneath the data points. The model predicts that, for low absolute currents, the difference is 0.44 of the modulation depth but a progressively smaller difference for higher absolute currents. The proportion is determined by the absolute current rather than the rate or the level in the perceptual dynamic range.

loudness perceived. Both model and experimental data show that amplitude modulation affects the loudness of a pulse train differently depending on the absolute levels used. The rate of stimulation and the loudness level of the stimulus only affect the way modulation changes loudness by way of influencing the absolute currents used. The results show that, for low currents, modulated stimuli with equal average currents have approximately equal loudness percepts. A rough rule of thumb from these data is that a ‘low’ absolute current would be 30–40 current steps or 5–7 dB below the comfortably loud level for a 500-Hz pulse train, which is generally achieved for overall rates more than 8 kHz. However, for higher absolute currents, the modulated stimulus will be louder than the constant-current pulse train of equal average current level.

Before discussing the implications of these findings for interpretation of previous research, it is important to discuss whether the findings would be generalisable to modulation frequencies lower than those used in the experiment and for modulation of phase duration rather than current.

Published literature about modulation detection reports investigations of a wider range of modulation rates than are investigated in our loudness experiment. Many papers have concentrated on modulations in the 20–50 Hz range that are associated with important speech envelope information, and others

have measured temporal modulation transfer functions with modulation rates up to 800 Hz. The loudness model used in the present paper assumes that loudness is related to the output of a temporal integration window (McKay and McDermott 1998) of several milliseconds. This output will be steady for modulation frequencies above about 200–300 Hz but will fluctuate at the modulation frequency for modulation frequencies below this, with progressively larger fluctuations as the modulation frequency decreases. Assuming that the fluctuations in integrator output are what are perceived as modulation in a modulation detection task, the model would then predict the well-known low-pass characteristic of temporal modulation transfer functions.

Although the loudness model does not specify how an overall stimulus loudness would be derived from a fluctuating integrator output, it is reasonable to assume that it would be derived from an average output [or average of ‘looks’ if the integration window can be considered a ‘look’ in a multiple-look model of loudness (Viemeister and Wakefield 1991)] unless the fluctuations were so slow (perhaps less than 2–3 Hz) that the stimulus would be described by the listener as having a changing loudness. The loudness model predicts no effect of modulation frequency on loudness for modulation frequencies greater than about 300 Hz. If fluctuations in the integrator output are averaged for overall stimulus loudness, then the

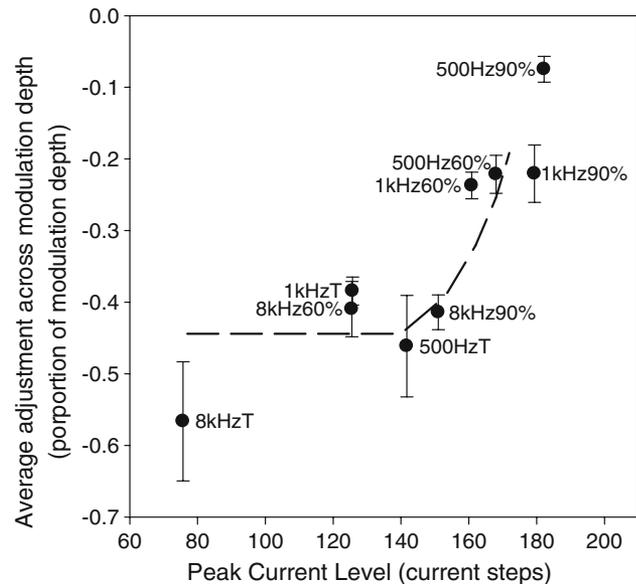


FIG. 4. Experimental data for the difference (averaged across modulation depth) between peak current in the modulated stimulus and the equally loud fixed current level, expressed as a proportion of the modulation depth. The *dashed line* represents the model data of the virtual subject from Figure 3 with a horizontal adjustment to account for differences in kneepoint in Eq. 1 (see text). Labels are to the left or right of the data points. The *error bars* represent standard errors.

predicted lack of effect of modulation frequency on overall loudness would also apply for modulation frequencies less than 300 Hz. In this case, the effect of modulation on loudness for modulations in the speech envelope range would be very similar to the effect at higher modulation frequencies. This proposition is supported by a preliminary work in our laboratory for a different experiment (Fraser and McKay, unpublished data) in which lower modulation frequencies have been balanced in loudness with unmodulated stimuli. The preliminary data show an analogous effect of modulation on loudness (i.e. modulation increases loudness if the average current is kept constant) to that shown in this experiment.

Some experiments using modulated stimuli have modulated the phase duration of the pulses rather than their current amplitude. Usually, a dB change of phase duration produces less loudness change than the same dB change of current amplitude, and the difference between the two is subject- and level-dependent (see McKay and McDermott 1999). However, with modulation detection tasks, the modulation amplitude is generally small enough to assume a linear, if not equal, relationship between phase duration and amplitude changes (on log scales), so the loudness model would apply in principle.

Implications for interpretation of previous experiments

The effect of modulation on overall loudness has implications both for the design and interpretation of psychophysical experiments using modulated stimuli. If current is modulated around a certain value and the carrier rate is lower than about 8 kHz, then the modulated stimulus will become louder as the modulation depth increases, with this effect increasing at higher current levels. In modulation detection tasks, therefore, it is important to balance the loudness of modulated with unmodulated stimuli and to use level jitter in order to minimise cues due to loudness differences between the stimuli. If overall loudness is not controlled, the threshold of modulation detection may be determined by the detection of overall loudness change rather than the detection of modulation per se, depending on which cue is the more salient for the subject for that particular stimulus. Even if the loudness and modulation cues are perceptually orthogonal, the combination of the two cues would lead to a more advantageous modulation detection threshold regardless of whether each cue alone is sub-threshold.

Published articles (Shannon 1992; Busby et al. 1993; Cazals et al. 1994; Chatterjee and Robert 2001; Fu 2002; Chatterjee and Oba 2005; Galvin and Fu 2005; Pfungst et al. 2007; Galvin and Fu 2009) that have explored the detection of modulation in coch-

lear implantees with pulse trains have generally used a sinusoidal current or phase duration modulation and a detection task in which the subject must discriminate the modulated stimulus from an unmodulated one in an adaptive n -interval forced choice task. Generally, the modulated stimuli were implicitly assumed to be of equal loudness to the unmodulated one, as no procedures were used to minimise overall loudness cues. Although Pfungst et al. (2007) reported that the modulated stimuli were louder than unmodulated stimuli based on lower average C levels, they dismissed the possibility of overall loudness cues on the basis that the variance in setting C levels by adjustment was greater than the average expected adjustment for equal loudness at modulation detection threshold. However, this argument is not valid, since the smallest current level difference detectable in an n -interval forced choice task (i.e. intensity DL) is far smaller than the variance of setting a C level by adjustment. Galvin and Fu (2009) also mentioned in their discussion that there might be overall loudness cues but did not specifically test this possibility or take action to minimise them when collecting the data.

In order to get a sense of the potential interference of loudness cues, Table 3 contains a conversion between the unit of modulation depth that is commonly used in modulation detection tasks [$20\log(m)$, where $m \times 100$ is the percentage of current modulation around a mean] and the percent modulation and the equivalent peak-to-peak modulation depth in the current step units used in this study. By reference to Figures 1 or 2, and the carrier rate of the stimuli, the overall loudness change caused by the modulation can be estimated in terms of predicted adjustment (in current step units) needed to make it the same

TABLE 3

Conversion table between the unit used commonly in modulation detection experiments [$-20\log(m)$], the per cent equivalent current change (relative to unmodulated current) and the equivalent modulation depth in current steps of 0.1759 dB

$-20\log(m)$	Per cent current modulation	Modulation depth (peak to peak) in current steps of 0.1759 dB
5	56	62
10	32	33
14	20	20
20	10	10
30	3.2	3
40	1	1
50	0.4	0.4

Using this conversion and the carrier rate of the stimulus, the reader can estimate from Figures 1 or 2 the loudness change of the modulated stimulus in units of predicted current-step adjustment in the unmodulated stimulus needed to balance the loudness

loudness as the unmodulated pulse train of the same average current. Taking the example of a 500-Hz carrier rate at a comfortably loud level, the current adjustment away from average current would be more than one current step by the time the modulation depth was three to six current steps. If we assume that a subject could detect the loudness change evoked by a one-current step increment, then any modulation detection thresholds greater than three to six current steps could be due to a loudness cue. This translates to $-20\log(m)$ values smaller than 30 dB. In contrast, near the threshold for the 500 Hz carrier, the difference between the average current and the equally loud current will not become obvious until the modulation depth is at least 20–30 current steps or $-20\log(m)$ smaller than 10–14 dB. Of course, the potential loudness cue relies on the intensity difference limens of the particular subject, which are also strongly inversely correlated with absolute current levels (Kreft et al. 2004). However, Nelson et al. (1996) showed that some subjects could indeed detect a one-current-step difference or smaller. For a 1-kHz carrier rate, the modulation depths with potential loudness cues range from more than five to ten current steps at higher levels [$-20\log(m)$ smaller than 25 dB] to more than 20–30 current steps near the threshold [$-20\log(m)$ values smaller than 10–14 dB].

Cazals et al. (1994) measured detection thresholds using a carrier rate of 1,667 Hz and modulation frequencies between 20 and 800 Hz at three different current levels. They found better detection at higher levels, and curiously, they found that subjects who did *not* exhibit a low-pass characteristic for modulation frequency had poorer speech perception. They suggested that detection of high frequency modulation might interfere with speech understanding. It can be observed that the modulation detection in each subject showed little change between 200 and 800 Hz and ranged from $-20\log(m)$ between 9 and 1 at the lowest level (modulation depth of 40–40 current steps equivalent), between 12 and 7 (25–50 current steps equivalent) at the mid-level (with the exception of one subject) and ranged from 25 to 12 at the highest level (6–25 current steps equivalent). These ranges correspond very well to those expected if subjects were using *solely loudness cues* at these higher modulation frequencies. Thus, a more satisfying explanation for the lack of low-pass characteristic in some subjects is that those subjects were never able to take advantage of a modulation cue for *lower* modulation frequencies and thus had the poorest temporal resolution and hence the poorest speech perception.

A second interesting study is that of Fu (2002), who found a correlation between the modulation detection thresholds averaged across different levels and speech understanding. He used a 1,000-Hz carrier

rate and modulation frequency of 20 Hz. If one considers a hypothetical subject who had no temporal resolution ability at all but relatively good intensity discrimination, one would predict that the modulation detection threshold [$-20\log(m)$] would range from below 14 near threshold to below 25 at the higher levels in the dynamic range based solely on loudness cues (as discussed above). Indeed there were two subjects (S8 and S9) who exhibited precisely this pattern. Another three subjects (S5, S6 and S7) showed the same pattern in low–mid-dynamic range and then improved their modulation detection to levels that make loudness cues unlikely. Only the ‘best’ four subjects showed modulation detection thresholds better than predicted by loudness cues alone over the whole dynamic range. It would be interesting to see what the results would look like if loudness cues were removed: ‘poor’ detection thresholds might look even poorer (as these are the ones most likely to be influenced by loudness in the current data), making the differences between subjects more stark than presented.

Galvin and Fu (2005), and expanded in Galvin and Fu (2009), compared modulation detection thresholds for carrier rates between 250 and 2,000 Hz, and modulation frequencies between 5 and 100 Hz. Similarly, Pfingst et al. (2007) measured modulation detection thresholds for carriers of 250 and 4,000 Hz carriers and modulation rate of 40 Hz. Both authors found better modulation detection thresholds at the lower carrier rates and concluded that the use of high carrier rates in speech processors was a disadvantage. The carrier rates chosen, however, would evoke very different loudness cues, with the 250 Hz carrier rate having significant loudness cues even for carrier levels near threshold. For the 250 Hz carrier at higher levels the peak level in the signal will essentially determine loudness (see Fig. 2), so any modulation even as small as $20\log(m)=-40$ (equivalent to one current step) might lead to perceptually significant loudness change. An inspection of Figure 1 in Galvin and Fu (2005) (comparing 250 and 2,000 Hz carrier rates) and Figure 2 in Pfingst et al. (2007) (comparing 250 and 4,000 Hz carriers) shows that very few if any data points for the 250 Hz carrier rate would be definitely safe from loudness cues. In Galvin and Fu (2009), only mean modulation thresholds are presented, but their level indicates that the difference between the carrier rates may be contributed by greater loudness cues at the lower rates. It must be concluded that any interpretation from these experiments of effect of carrier rate on temporal processing ability cannot be safe. As an aside, it should be pointed out that, in all three papers, the different dynamic ranges of the carrier rates was not taken into account when assessing the practical relevance of the result. As the

dynamic ranges of the 2,000 or 4,000 Hz stimuli were about double those of the 250 Hz stimuli, any modulation in the acoustic signal would be translated into a larger modulation in the electrical signal by the speech processor running at 2,000 or 4,000 Hz, compared to the one running at 250 Hz, ameliorating or potentially reversing any potential effects of carrier rate on modulation detection when expressed in $20\log(m)$. If the effect of potentially greater loudness cues at the lower rates is also considered, the possibility of a conclusion opposite to those stated arises (i.e. higher rates might facilitate better temporal resolution). This possibility however but would have to be tested in a different experiment.

Next, it is interesting to discuss the possible influence of loudness cues when random ‘noise’ is added to modulation. Chatterjee and Robert (2001) and Chatterjee and Oba (2005) investigated the effect of adding noise (by adding a random amount of current to individual pulses) on modulation detection thresholds. They used carrier rates of 500 or 1,000 Hz in both papers and phase duration modulation in the former paper and current amplitude modulation in the latter paper. There are two main aspects of the reported results that would suggest that cues other than modulation were influencing results in some conditions. First, Chatterjee and Robert (2001) showed that noise effectively flattened the modulation transfer function for most subjects between modulation frequencies of 25 and 300 Hz (Fig. 5 in their paper). Moreover, it can be seen from that figure that the modulation detection thresholds in the greatest amounts of noise (between 20% and 40%) are very similar across subjects and modulation frequencies [$20\log(m) = -20$ or 10% approximately] regardless of the differences in modulation detection ability in quiet. This is a strong indication that, in the highest noise conditions, modulation detection was dominated by something other than temporal resolution.

Secondly, in both papers, it was found that noise improved modulation detection *only* in conditions where the modulation detection in quiet was very poor and *essentially always* decreased detection ability for very high noise levels (20–40%), leading to a non-monotonic function for those conditions with poor detection in quiet. This non-monotonicity was attributed to a possible stochastic resonance effect, whereby there was an optimum amount of noise for modulation detection. However, it is precisely these conditions (greater modulation depth) in which loudness cues are more likely to dominate over temporal processing cues. The loudness model could provide a possible explanation for why the noise had a non-monotonic effect on modulation detection for conditions with poor detection threshold in quiet. If one considers a hypothetical subject with no temporal

processing ability who was doing the task via loudness cues alone, there are two counteracting effects that could lead to this non-monotonic behaviour. First, if one considers the noise as providing an extra modulation alone (in terms of variation in current levels) so that the task in noise was a discrimination of modulation amplitude change task rather than a modulation detection task, the model would predict that the loudness cue would be more salient as the noise increased. It can be seen from Figure 1 that the predicted effect of modulation depth on loudness is non-linear: increasing the modulation depth from five to ten current steps increases the mismatch between predicted and average current by a greater amount than increasing it from zero to five steps, etc. Therefore, if the noise were merely increasing the peak-to-peak variation in the signal, the model would predict a continually improving signal detection (based on

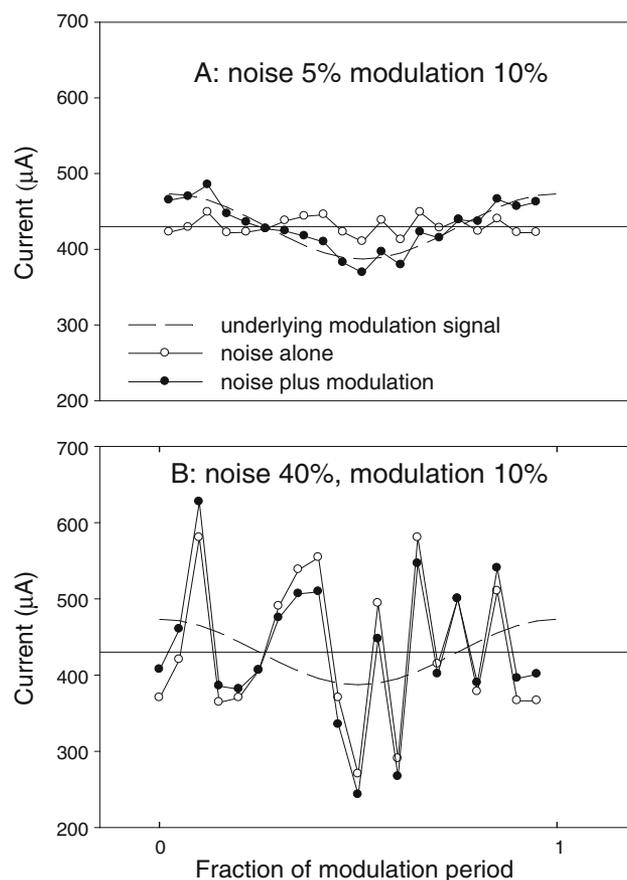


FIG. 5. Example graphs showing the current values for sequential pulses in unmodulated and modulated stimuli over one period of modulation. The underlying modulation is represented by the *dashed line* in each panel. Current values for the noise-only stimulus are represented by *open symbols* and those for the noise plus modulation by *closed symbols*. The *horizontal line* is the average current value. The same representative noise (but scaled differently) is used in each panel. In **A**, the noise is smaller than the modulation, and in **B**, it is bigger. In **A**, the modulation determines which pulses have the values most different from the mean, and in **B**, the noise determines which values are more different from the mean.

loudness cues) with increasing 'noise'. However, for large amounts of noise (relative to the modulation), the 'extreme' current values (furthest from the mean) are determined more by the noise than the added modulation, decreasing the amount of loudness cue due to the modulation. Figure 5 illustrates the different impact of added modulation for two different noise levels. When the noise is smaller than the modulation (subpanel A), the modulation essentially just increases the modulation depth, and the extreme current values occur at the extremes of the modulation; when the noise is larger than the modulation (subpanel B), the extreme values in the modulated stimulus are more determined by the noise than the modulation. Inspecting the data from the two published papers in conditions for which the modulation detection in quiet was very poor (generally when higher modulation rates or lower current levels were used), it can be seen that the position of the 'optimum' noise is dependent on the size of the modulation signal in quiet: The greater the modulation of the signal needed in quiet, the more noise is needed to reach 'optimum' noise level.

The above considerations lead to the conclusion that the effect attributed to stochastic resonance in these papers is more likely to be due to the influence of loudness cues in the detection task.

CONCLUSIONS

In this paper, we have demonstrated that the modulation of a pulse train affects the loudness perceived in a way that depends upon the absolute current levels used, as predicted by the loudness model of McKay et al. (2003) and not on the perceptual level of the stimulus. Furthermore, we have demonstrated that the effect of modulation on loudness should be considered by researchers and steps taken to avoid loudness cues when their goal is to determine the temporal processing ability of implantees via modulation detection tasks. This is especially important for conditions where the modulation detection is poorer (requiring use of larger modulation depths that could potentially lead to perceptible loudness changes) and for stimuli that use higher currents on an absolute scale. We have shown that the results of previous experiments measuring modulation detection have led to some inferences about signal processing that can be considered unsafe if loudness cues are considered. In particular, the conclusions that lower rates (compared to higher rates) improve modulation detection or that adding random noise fluctuations to

a signal could provide better transfer of temporal information should be re-assessed in other studies that control loudness cues.

ACKNOWLEDGEMENTS

This research was financially supported by the Australian National Health and Medical Research Council and the Garnet Passe and Rodney Williams Memorial Foundation. We gratefully thank the cochlear implantees who freely gave their time to participate in these experiments. We thank Chris Plack and Hugh McDermott for helpful comments on a draft manuscript.

REFERENCES

- BUSBY PA, TONG YC, CLARK GM (1993) The perception of temporal modulations by cochlear implant patients. *J Acoust Soc Am* 94:124–131
- CAZALS Y, PELIZZONE M, SAUDAN O, BOEX C (1994) Low-pass filtering in amplitude modulation detection associated with vowel and consonant identification in subjects with cochlear implants. *J Acoust Soc Am* 96:2048–2054
- CHATTERJEE M (1999) Effects of stimulation mode on threshold and loudness growth in multielectrode cochlear implants. *J Acoust Soc Am* 105:850–860
- CHATTERJEE M, OBA SI (2005) Noise improves modulation detection by cochlear implant listeners at moderate carrier levels. *J Acoust Soc Am* 118:993–1002
- CHATTERJEE M, ROBERT ME (2001) Noise enhances modulation sensitivity in cochlear implant listeners: stochastic resonance in a prosthetic sensory system? *J Assoc Res Otolaryngol* 2:159–171
- FU QJ (2002) Temporal processing and speech recognition in cochlear implant users. *Neuroreport* 13:1635–1639
- GALVIN JJ 3RD, FU QJ (2005) Effects of stimulation rate, mode and level on modulation detection by cochlear implant users. *J Assoc Res Otolaryngol* 6:269–279
- GALVIN JJ 3RD, FU QJ (2009) Influence of stimulation rate and loudness growth on modulation detection and intensity discrimination in cochlear implant users. *Hear Res* 250:46–54
- KREFT HA, DONALDSON GS, NELSON DA (2004) Effects of pulse rate and electrode array design on intensity discrimination in cochlear implant users. *J Acoust Soc Am* 116:2258–2268
- MCDERMOTT HJ, MCKAY CM, RICHARDSON LM, HENSHALL KR (2003) Application of loudness models to sound processing for cochlear implants. *J Acoust Soc Am* 114:2190–2197
- MCDERMOTT HJ, SUGHER CM, MCKAY CM (2004) Speech perception with a cochlear implant sound processor incorporating loudness models. *Acoust Res Lett Online* 6:7–13
- MCKAY CM, MCDERMOTT HJ (1998) Loudness perception with pulsatile electrical stimulation: the effect of interpulse intervals. *J Acoust Soc Am* 104:1061–1074
- MCKAY CM, MCDERMOTT HJ (1999) The perceptual effects of current pulse duration in electrical stimulation of the auditory nerve. *J Acoust Soc Am* 106:998–1009
- MCKAY CM, REMINE MD, MCDERMOTT HJ (2001) Loudness summation for pulsatile electrical stimulation of the cochlea: effects of rate, electrode separation, level, and mode of stimulation. *J Acoust Soc Am* 110:1514–1524

- MCKAY CM, HENSHALL KR, FARRELL RJ, MCDERMOTT HJ (2003) A practical method of predicting the loudness of complex electrical stimuli. *J Acoust Soc Am* 113:2054–2063
- NELSON DA, SCHMITZ JL, DONALDSON GS, VIEMEISTER NF, JAVEL E (1996) Intensity discrimination as a function of stimulus level with electric stimulation. *J Acoust Soc Am* 100:2393–2414
- PFGST BE, XU L, THOMPSON CS (2007) Effects of carrier pulse rate and stimulation site on modulation detection by subjects with cochlear implants. *J Acoust Soc Am* 121:2236–2246
- SHANNON RV (1985) Threshold and loudness functions for pulsatile stimulation of cochlear implants. *Hear Res* 18:135–143
- SHANNON RV (1992) Temporal modulation transfer functions in patients with cochlear implants. *J Acoust Soc Am* 91:2156–2164
- VIEMEISTER NF, WAKEFIELD GH (1991) Temporal integration and multiple looks. *J Acoust Soc Am* 90:858–865
- ZENG FG, SHANNON RV (1994) Loudness-coding mechanisms inferred from electric stimulation of the human auditory system. *Science* 264:564–566
- ZENG FG, SHANNON RV (1995) Loudness of simple and complex stimuli in electric hearing. *Ann Otol Rhinol Laryngol Suppl* 166:235–238
- ZHANG C, ZENG FG (1997) Loudness of dynamic stimuli in acoustic and electric hearing. *J Acoust Soc Am* 102:2925–2934